

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 14:50:12 ; Search time 18 Seconds  
(without alignments)  
809.904 Million cell updates/sec

Title: US-09-831-805A-6

Perfect score: 1635  
Sequence: 1 MALRRPRLRLCARLPDFL.....VNYINTDEGFRHKSSPVI 310

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	ID	Description
1	483	29.5	298	1 JAM2 HUMAN
2	449.5	27.5	300	1 JAM1 MOUSE
3	419	25.6	299	1 JAM1 HUMAN
4	403.5	24.7	298	1 JAM1 BOVIN
5	231.5	14.2	319	1 A33 HUMAN
6	186	11.4	1088	1 NCAL XENLA
7	180	11.0	333	1 ANML DROME
8	177	10.8	1092	1 NCAL XENLA
9	171.5	10.5	725	1 NCAL MOUSE
10	171.5	10.5	1115	1 NCAL MOUSE
11	171	10.5	1377	1 NEOL RAT
12	170.5	10.4	858	1 NCAL RAT
13	169.5	10.4	1912	1 PFPD HUMAN
14	165	10.1	837	1 NCAL HUMAN
15	164.5	10.1	1051	1 PTK7 CHICK
16	164.5	10.1	3707	1 PGBM MOUSE
17	164	10.0	837	1 NCAL MOUSE
18	163.5	10.0	853	1 NCAL BOVIN
19	163	10.0	1493	1 NEOL MOUSE
20	162.5	9.9	6632	1 UN89 CAEEL
21	161.5	9.9	365	1 CXAR HUMAN
22	158	9.7	1091	1 NCAL CHICK
23	155.5	9.5	761	1 NCAL HUMAN
24	155.5	9.5	848	1 NCAL HUMAN
25	155.5	9.5	1284	1 NRCA CHICK
26	153.5	9.4	1461	1 NEOL HUMAN
27	153	9.4	344	1 NTRI HUMAN
28	151.5	9.3	349	1 LACH SCHAM
29	151	9.2	353	1 CBPU CHICK
30	150.5	9.2	702	1 CEAS HUMAN
31	149.5	9.1	1897	1 PTFP HUMAN
32	149	9.1	338	1 LAMP CHICK
33	148.5	9.1	338	1 LAMP RAT

## RESULT 1

JAM2 HUMAN	148.5	9.1	646	1	MU18_HUMAN
AC	148	9.1	344	1	NTRI_MOUSE
AC	147	9.0	344	1	NTRI RAT
37	146.5	9.0	338	1	LAMP_HUMAN
38	146	8.9	365	1	CXAR_MOUSE
39	145.5	8.9	345	1	OPCM_RAT
40	145.5	8.9	4391	1	PGBM_HUMAN
41	145	8.9	521	1	CEAL_MOUSE
42	144	8.8	847	1	CD22_HUMAN
43	143.5	8.8	1443	1	NEOL_CHICK
44	141	8.6	345	1	OPCM_BOVIN
45	139	8.5	1040	1	AXOL_RAT

P43121 homo sapien  
 Q99pf0 mus musculus  
 Q62718 rattus norv  
 Q13449 homo sapien  
 P97792 mus musculus  
 P32736 rattus norv  
 P98160 homo sapien  
 P31809 mus musculus  
 P20273 homo sapien  
 Q90610 gallus gall  
 P11834 bos taurus  
 P22063 rattus norv

JAM2 HUMAN STANDARD; PRT; 298 AA.  
 AC P57087;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Junctional adhesion molecule 2 precursor (Vascular endothelial  
 DE junction-associated molecule) (VE-JAM).  
 GN JAM2 OR VEJAM OR C21ORP43.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Vascular endothelial cells;  
 RX MEDLINE=20317114; PubMed=10779521;  
 RA Palmeri D., van Zante A., Huang C.C., Hemmerich S., Rosen S.D.;  
 RT "Vascular endothelial junction-associated molecule, a novel member of  
 RT the immunoglobulin superfamily, is localized to intercellular  
 RT boundaries of endothelial cells";  
 RL J. Biol. Chem. 275:19139-19145(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=20507930; PubMed=10945976;  
 RA Cunningham S.A., Arrate M.P., Rodriguez J.M., Bjerkce R.J.,  
 RA Vanderslice P., Morris A.P., Brock T.A.;  
 RT "A novel protein with homology to the junctional adhesion molecule:  
 RT Characterization of leukocyte interactions";  
 RL J. Biol. Chem. 275:34750-34756(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=2398257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Faney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

## ALIGNMENTS



```

DR MGD; MGI:1321398; Flr1.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00406; Ig; 1.
DR PROSITE; PS00835; IG LIKE; 2.
KW Tight junction; Immunoglobulin domain; Glycoprotein; Transmembrane;
KW Repeat; Signal; 3D-structure.
FT SIGNAL 1 26
FT CHAIN 27 300 JUNCTIONAL ADHESION MOLECULE 1.
FT DOMAIN 27 238 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 239 259 POTENTIAL.
FT DOMAIN 260 299 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 28 122 IG-LIKE V-TYPE 1.
FT DOMAIN 134 230 IG-LIKE V-TYPE 2.
FT DISULFID 49 108 POTENTIAL.
FT DISULFID 152 212 POTENTIAL.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 300 AA; 32368 MW; 391F3E48FE3B97EC CRC64;

Query Match 27.5%; Score 449.5; DB 1; Length 300;
Best Local Similarity 35.2%; Pred. No. 1.1e-30;
Matches 105; Conservative 56; Mismatches 122; Indels 15; Gaps 7;

QY 20 LLLFRGCLIGAV-----NLKSSNTPVVPQFSEVELSCIITDSQTSPPRIEKKIODEQ 74
DB 11 LFLFTSMILGSLVGKGSVYTAQDVQVPENESIKLTC--TYSGFSSPRVKEKVGQST 68
QY 75 TTYVPFDNKIQGLAGRAEILGKTSKLTWNVTRDSALYRCVEVARNDKDEIVIELT 134
DB 69 TALVCYNQITAPYADRV-TFSSSGITFSSVTRKNGEYTC-MVSEEGQNGYGVSIHLT 126
QY 135 VQVKVPTVCPVKAPVKGKATLHCQSEGHPRPHYSWYRNDVPLPT-DSRANRPRFNS 193
DB 127 VLVPSPKPTISVPSSVITGNRAVLTCSEHDGSPSEYSNFKDGISMLTADAKTKRAFMS 186
QY 194 SSSLNSETGLVFTAVHKDDSQYICIASNDAGSA-RCEQOMEVYDNLNIGIIGVLVW 252
DB 187 SFTIDPKSGDLIFDPTAFDSEYVQQAQNGVGTAMRSEAAHMDAVELNVLGVIIVAVLVT 246
QY 253 LAVLALITLIGCCAYRRGYFINNKQDGSYKPNPGKPDGVNVRTDEEDGRHKSSFVI 310
DB 247 LILLGLLIFGVWFAYSRGVFETTKG-----TAPGKKVIYSQPSISEGEFKTSSEFLV 300

RESULT 3
ID JAM1 HUMAN STANDARD; PRT; 299 AA.
AC Q9Y6Z4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Juncional adhesion molecule 1 precursor (JAM) (Platelet adhesion
DE molecule 1) (PAM-1) (Platelet F11 receptor).
GN F11R OR JAM1 OR JCAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99323940; PubMed=10395639;
RA Ozaki H., Ishii K., Horiuchi H., Arai H., Kawamoto T., Okawa K.,
RA Iwamatsu A., Kita T.;
RT "Combined treatment of TNF-alpha and IFN-gamma causes redistribution
RT of junctional adhesion molecule in human endothelial cells.";
RL J. Immunol. 163:553-557(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Sobocka M.B., Sobocki T., Rushbrook J.I., Banerjee P., Weiss C.,
RA Kornecki E.;

```

```

RT "Molecular cloning and sequencing of the cDNA of F11 receptor, a
RT novel Ig superfamily member from human platelets.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Naik U.P., Naik M.U., DeLeon P., Spychala J.;
RT "Cloning and characterization of PAM-1, a novel platelet adhesion
RT molecule involved in platelet activation.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansorge W., Bloecher M., Bloecher H., Bauersachs S., Blum H.,
RA Lauber J., Dueterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
CC -!- FUNCTION: Seems to plays a role in epithelial tight junction
CC formation. Appears early in primordial forms of cell junctions and
CC recruits PARD3. The association of the PARD6-PARD3 complex may
CC prevent the interaction of PARD3 with JAM1, thereby preventing
CC tight junction assembly (By similarity). Plays a role in
CC regulating monocyte transmigration involved in integrity of
CC epithelial barrier. Involved in platelet activation.
CC -!- SUBUNIT: Interacts with the first PDZ domain of PARD3. The
CC association between PARD3 and PARD6B probably disrupts this
CC interaction (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Localized at tight junctions of both
CC epithelial and endothelial cells.
CC -!- SIMILARITY: BELONGS TO THE IMMUGLOBULIN SUPERFAMILY.
CC -!- SIMILARITY: Contains 2 immunoglobulin-like V-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AF111713; AAD42050.1; -
CC EMBL; AF207907; AAF22829.1; -
CC EMBL; AF172398; AAD48877.1; -
CC EMBL; AL136649; CAB66584.1; -
CC PIR; A59406; S56749.
CC Genew; HGNC:14685; F11R.
CC MIM; 605721; -
CC GO; GO:0006954; P:inflammatory response; TAS.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003006; Ig MHC.
CC InterPro; IPR003596; Ig_v.
CC Pfam; PF00047; Ig; 2.
CC SMART; SM00406; Ig; 1.
CC PROSITE; PS00835; IG LIKE; 2.
KW Tight junction; Immunoglobulin domain; Glycoprotein; Transmembrane;
KW Repeat; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 299 JUNCTIONAL ADHESION MOLECULE 1.
FT DOMAIN 26 238 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 239 259 POTENTIAL.
FT DOMAIN 260 299 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 27 125 IG-LIKE V-TYPE 1.
FT DOMAIN 135 228 IG-LIKE V-TYPE 2.
FT DISULFID 50 109 POTENTIAL.
FT DISULFID 153 212 POTENTIAL.
FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 299 AA; 32583 MW; D95DE2FEA23D2851 CRC64;

```

```
Query Match      25.6%; Score 419; DB 1; Length 299;
Best Local Similarity 32.8%; Pred. No. 4e-28;
Matches 101; Conservative 47; Mismatches 138; Indels 22; Gaps 7;

QY 8 RRLCARLPDFFLLFRGLGAVNLKSSNRTPVQVFESVELSCIITDQSDPRIEW 67
DB 9 RKLLCL----FILAILLCSLALSVTHSSFEVRIPEPNPVKLSLAY--SGFSSPRV 62
QY 68 KKIQDEQTYVFFDNKIQDLAGRAEILGKTSLKWNVTRDLSALYRCEVVARNDK 127
DB 63 KPDQGTTRLVCNNKINTASVEDRTVFL-PTGITFKSVTRDGTGYTC-MVSEEGNSY 120
QY 128 EIVIELTVQKVPVPCVRKAPVGVKMATLHCQSEGHPRPHYSWYRNDVPLPTDS 187
DB 121 EVKVKLVLPSPKPTNVPSSATIGNRAVLTCSEQDQSPPEYTWPKDGIWPTNP 180
QY 188 PRFNSSHLNSETGLVFTAVHKDDSGQYCYCIASNDAGSARCEBQ-EMEYVD 246
DB 181 RAFNSSYVLNPTTGLVDFPLSDATGEYSCEARNGYGTPTMSNAVRMEAVR 240
QY 247 GGVLVAVLALITLIGICAVRRGYFINNKODGES----YKNPKPDGVNVRTD 302
DB 241 AAVLVTLILLGLVFGIWFAYSRGHFDRTKGTSSKKVIYOPS-----AR 291
QY 303 RHKSSFVI 310
DB 292 KQTSSFLV 299

RESULT 4
JAM1_BOVIN
ID JAM1_BOVIN STANDARD; PRT; 298 AA.
AC Q9XT56.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Junctional adhesion molecule 1 precursor (JAM).
GN F11R OR JAM1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99323940; PubMed=10395639;
RA Ozaki H., Ishii K., Horiuchi H., Arai H., Kawamoto T., Okawa K.,
RA Iwamatsu A., Kita T.;
RT "Combined treatment of TNF-alpha and IFN-gamma causes redistribution
RT of junctional adhesion molecule in human endothelial cells.";
RL J. Immunol. 163:553-557(1999).
CC -!- FUNCTION: Seems to play a role in epithelial tight junction
CC formation. Appears early in primordial forms of cell junctions and
CC recruits PAR3. The association of the PAR6-PARD3 complex may
CC prevent the interaction of PAR3 with JAM1, thereby preventing
CC tight junction assembly (By similarity). Plays a role in
CC regulating monocyte transmigration involved in integrity of
CC epithelial barrier. Involved in platelet activation.
CC -!- SUBUNIT: Interacts with the first PDZ domain of PAR3. The
CC association between PAR3 and PAR6B probably disrupts this
CC interaction (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Localized at tight junctions of both
CC epithelial and endothelial cells.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -!- SIMILARITY: Contains 2 immunoglobulin-like V-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC

or send an email to license@isb-sib.ch.
-----
CC EMBL; AF111714; AAD42051.1; -.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_C2.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00047; Ig_2.
CC SMART; SM00408; IGC2; 1.
CC PROSITE; PS00835; IG LIKE; 2.
CC Tight junction; Immunoglobulin domain; Glycoprotein; Transmembrane;
CC Repeat; Signal.
CC SIGNAL 1 24 POTENTIAL.
CC FT CHAIN 25 298 JUNCTIONAL ADHESION MOLECULE 1.
CC FT DOMAIN 25 237 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 238 258 POTENTIAL.
CC FT DOMAIN 259 298 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 28 124 IG-LIKE V-TYPE 1.
CC FT DOMAIN 134 227 IG-LIKE V-TYPE 2.
CC FT DISULFID 49 108 POTENTIAL.
CC FT DISULFID 152 211 POTENTIAL.
CC FT CARBOHYD 184 184 N-LINKED (GLCNAC...) (POTENTIAL).
CC SQ SEQUENCE 298 AA; 32456 MW; 714FE1C1714769A2 CRC64;

Query Match      24.7%; Score 403.5; DB 1; Length 298;
Best Local Similarity 33.8%; Pred. No. 8.2e-27;
Matches 102; Conservative 48; Mismatches 125; Indels 27; Gaps 10;

QY 21 LLLFRGCLII-----GAVNLKSSNRTPVQ--EFESVELSCIITDQSDPRIEWKKIQ 71
DB 12 LLLFTSMILSLALGRGAV----QTYEPVVRVPPNNPAKLSG--SYSGFSPRVWEKETH 65
QY 72 DEQTYVFFDNKIQDLAGRAEILGKTSLKWNVTRDLSALYRCEVVARNDKDEIVI 131
DB 66 GDIRGLVCYNNKITASYENRV-TFSDTGITHSVTRKDTGMVTC-MVSEGGNTYGEVTV 123
QY 132 ELTVQKVPVTPVCRKAPVGVKMATLHCQSEGHPRPHYSWYRNDVPLPTDSRANPR 191
DB 124 QLIVLVPPSKPTINVPSSVTIGTRAVLTCSERDGSPPEYKWKFDGVEMPLEKSNRAPS 183
QY 192 NSSHLNSETGLVFTAVHKDDSGQYCYCIASND-AGSARCEBQEMEYVDNLIGLIGVL 250
DB 184 NSSVTLNQTGELIFDPVPSASDTGDTGCOQNGVSPVKSVDTVHMDAVELNVGGIAV 243
QY 251 VVLAVLALITLIGICAVRRGYFINNKODGESYKPKPGDGVNVRTD--EEGDFRHKSSF 308
DB 244 VTLILLGLALIFGIWFAYSRGIFDRAKK-GTSNKK-----VIYQPNARSDEGFEQTSSF 296
QY 309 VI 310
DB 297 LV 298

RESULT 5
A33_HUMAN
ID A33_HUMAN STANDARD; PRT; 319 AA.
AC Q99795;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cell surface A33 antigen precursor (Glycoprotein A33).
GN GPA33.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Colon carcinoma;
EX MEDLINE=97165045; PubMed=9012807;
RA Heath J.K., White S.J., Johnstone C.N., Catimel B., Simpson R.J.,
RA Moritz R.L., Tu G.-F., Ji H., Whitehead R.H., Groenen L.C.,
RA Scott A.M., Ritter G., Cohen L., Welt S., Old L.J., Nice E.C.,
RA Burgess A.W.;
```



```

DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00047; ig; 5.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00408; IGC2; 5.
DR PROSITE; PS50835; IG-LIKE; 5.
KW Cell adhesion; Glycoprotein; Transmembrane; Repeat; Brain;
KW Immunoglobulin domain; Alternative splicing; Signal.
FT SIGNAL 1 19
FT CHAIN 20 1088
FT FT
FT FT
FT DOMAIN 20 705
FT TRANSMEM 706 723
FT DOMAIN 724 1088
FT DOMAIN 724 1088
FT DOMAIN 20 108
FT DOMAIN 113 202
FT DOMAIN 209 294
FT DOMAIN 303 397
FT DOMAIN 400 484
FT DOMAIN 512 589
FT DOMAIN 618 686
FT DOMAIN 149 153
FT DOMAIN 158 162
FT DISULFID 41 93
FT DISULFID 136 186
FT DISULFID 232 282
FT DISULFID 323 379
FT DISULFID 420 473
FT CARBOHYD 82 82
FT CARBOHYD 219 219
FT CARBOHYD 310 310
FT CARBOHYD 341 341
FT CARBOHYD 417 417
FT CARBOHYD 443 443
FT CARBOHYD 472 472
FT VARSPIC 804 1049
FT SQ
SQ SEQUENCE 1088 AA; 117778 MW; 62738B5B03F3B83 CRC64;
Query Match 11.4%; Score 186; DB 1; Length 1088;
Best Local Similarity 29.9%; Pred. No. 98-08; 82; Gaps 12;
Matches 63; Conservative 30; Mismatches 82; Indels 36;
QY 30 GAVNLK-----SSNRTPVQPEFSEVLSCTIIDSQDPF-IEWK-KIQDEQTVVFFDN 82
Db 105 GTVNLKIYQKLTFFKNAPTQPEFKEGEDAVIICDVSSIPSITWRHKGK-----VIFKK 159
QY 83 KIQGLAGRAEILGKTSKIMNVRDRDSALYRCE--VVARND--RKEIDEIV-IELTVQV 137
Db 160 DV-----RFVVLANNYLQIRGKTKTDEGTYRCEGRILARGEINKYDIQVIVNVPPTIQA 213
QY 138 KPTVPVCRVPKAVPVGKMATLHCQSEGHPRPHYSWYRNDVPLPTDSRANPRFRNSSHL 197
Db 214 RQL-----RVNATANMAESVVLSC-DADGFPDPPEISWLKKGEPI-EDGEEKISF----- 260
QY 198 NSETGLVFTAVHKDDSGGYVCIANSAGSA 228
Db 261 NEDQSEMTIHVVEKDDAEYSIANNQAGEA 291
RESULT 7
AMAL DROME
ID AMAL DROME STANDARD; PRT; 333 AA.
AC P15364; O9V3A5;
DT 01-APR-1990 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Almagam protein precursor.
GN AMA OR-BG:DS00276.6 OR CG2198.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R;
RX MEDLINE=89028670; PubMed=3141062;
RA Seger M.A., Haffley L., Kaufman T.C.;
RT "Characterization of analgam: a member of the immunoglobulin
RL Cell 55:589-600 (1988).
RN RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Celniker S.E., Pfeiffer B.D., Knafels J., Martin C.H., Mayeda C.A.,
RA Palazzolo M.J.;
RT "Complete sequence of the Antennapedia complex of Drosophila.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBAJ databases.
RN RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazew R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöckel A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei Y., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M.T., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN RP SEQUENCE FROM N.A.
RC STRAIN=Berkley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Kronmiller B., Pacle J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH080.1-RESEARCH080.8 (2002).
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
(Possible).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```





```
QY 30 GAVNLKSSNR-----TPVQVFESVELSCIITSDSPR-IEWK-KIQDQTTVVFFDN 82
D 105 GTVNLKIQKLTFFKYPPTQFTEGEDAVIICDVSSSIPSIITWRHKGKD-----VIFKK 159
QY 83 KIQDLAGRAELGLKTSKINVTFRDSALYRCE--VVARNDKREIDIVIELTVQVQKPV 140
D 160 DV-----RFVVLANNYLQIRGIKKTGNYRCEGRILARG---EINYKDIQVIVNVPPL 210
QY 141 TPV--CRYPKAVPVCKMATLHQESEGHPRHYSWYRNDVPLPTDSRANPRFRSSSHLN 198
D 211 IQARQIRVNANMDESIVLSC-DADGPDPEISWLKKGKEPT-EDGEBKISF-----N 261
QY 199 SETGLVFTAVHKDPSGOYICIASNDAGSA 228
D 262 EDKSEMTYRVEKEDAEYSCIANNQAGEA 291

RESULT 9
NCA2 MOUSE
ID _NCA2 MOUSE STANDARD; PRT; 725 AA.
AC P13594; Q61950;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neural cell adhesion molecule 1, 120 kDa isoform precursor (N-CAM 120)
DN (NCAM-120).
GE NCAM1 OR NCAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=87246524; PubMed=3595563;
RA Barthele D., Santoni M.-J., Wille W., Ruppert C., Caix J.-C.,
RA Hirsch M.-R., Fontecilla-Camps J.-C., Goridis C.;
RT "Isolation and nucleotide sequence of mouse NCAM cDNA that codes for
RL a Mr 79,000 polypeptide without a membrane-spanning region.";
RN [2]
RP SEQUENCE OF 20-700 FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=89251563; PubMed=2721486;
RA Santoni M.J., Barthele D., Vopper G., Boned A., Goridis C., Wille M.;
RT "Differential exon usage involving an unusual splicing mechanism
RL generates at least eight types of NCAM cDNA in mouse brain.";
RN [3]
RP SEQUENCE OF 642-725 FROM N.A.
RX MEDLINE=88283628; PubMed=3396534;
RA Barbas J.A., Chaix J.C., Steinmetz M., Goridis C.;
RT "Differential splicing and alternative polyadenylation generates
RL distinct NCAM transcripts and proteins in the mouse.";
RN [4]
RP SEQUENCE OF 20-36.
RX MEDLINE=86140120; PubMed=3512556;
RA Rougon G., Marshak D.R.;
RT "Structural and immunological characterization of the amino-terminal
RL domain of mammalian neural cell adhesion molecules.";
RL J. Biol. Chem. 261:3396-3401(1986).
CC -!- FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE INVOLVED IN
CC NEURON-NEURON ADHESION, NEURITE FASCICULATION, OUTGROWTH OF
CC NEURITES, ETC.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- ALTERNATIVE PRODUCTS:
CC Name=N-CAM 120; Sequence=Displayed;
CC Name=N-CAM 180;
CC Name=N-CAM 140;
CC Name=N-CAM 140;
```

```
CC CC IsoId=P13595-2; Sequence=External;
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 2 fibronectin type III domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y00051; CAA68263.1; -
CC EMBL; X15049; CAA33148.1; ALT_SEQ.
CC EMBL; X07195; CAA30173.1; -
CC PIR; A29673; IJMSNG.
CC PDB; 2NCM; 12-MAR-97.
CC MGD; MGI:97281; Ncam1.
CC InterPro; IPR003961; FN III.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_C2.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00041; fn3; 2.
CC Pfam; PF00047; Ig; 5.
CC SMART; SM00060; FN3; 2.
CC SMART; SM00408; IGC2; 5.
CC PROSITE; PS00835; Ig_Like; 5.
CC Cell adhesion; Glycoprotein; Repeat; Alternative splicing;
CC Immunoglobulin domain; Signal; Heparin-binding; GPI-anchor;
CC 3D-structure.
CC SIGNAL 1 19
CC CHAIN 20 725
CC
CC DOMAIN 20 111
CC DOMAIN 116 205
CC DOMAIN 212 302
CC DOMAIN 309 402
CC DOMAIN 407 492
CC DOMAIN 519 596
CC DOMAIN 152 625
CC DOMAIN 161 165
CC DISULFID 41 96
CC DISULFID 139 189
CC DISULFID 235 288
CC DISULFID 330 386
CC DISULFID 427 480
CC CARBOHYD 222 222
CC CARBOHYD 316 316
CC CARBOHYD 348 348
CC CARBOHYD 424 424
CC CARBOHYD 450 450
CC CARBOHYD 479 479
CC CONFLICT 261 268
CC CONFLICT 273 273
CC CONFLICT 354 355
CC CONFLICT 549 549
CC CONFLICT 572 572
CC CONFLICT 575 575
CC CONFLICT 589 594
CC CONFLICT 600 602
CC CONFLICT 657 657
CC SEQUENCE 725 AA; 80296 MW; C2AE8B84461C6B2F CRC64;
CC
CC Query Match 10.5%; Score 171.5; DB 1; Length 725;
CC Best Local Similarity 27.3%; Pred. No. 9.1e-07;
CC Matches 63; Conservative 41; Mismatches 84; Indels 43; Gaps 12;
CC
CC QY 27 CLIGA-----VNLKSSNR-----TPVQVFESVELSCIITSDSPR-DPIRWKKIQ 71
CC D 96 CVWTAEDGTQSEATVNVKIFQKLMFNAPTQEFKEGSDAVIVCDVSSLPPTTIWK--- 152
```





```
FT CARBOHYD 479 479 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 810 1076 Missing (in isoform N-CAM 140).
FT /FTIG=VSP 002588.
SQ SEQUENCE 1115 AA; 119351 MW; 2C93D0D474CFBCAF CRC64;

Query Match 10.5%; Score 171.5; DB 1; Length 1115;
Best Local Similarity 27.3%; Pred. No. 1.6e-06;
Matches 63; Conservative 41; Mismatches 84; Indels 43; Gaps 12;

QY 27 CLIGA-----VNLKSSNR-----TPVQEPESVLSCLITDSQTS-DPIRWKKIQ 71
DQ 96 CVVTAEDGTQSEATVNVKIFQKMPKNAPTPOEFKEGEDAVIVCDVSSLPPTIWK--- 152
QY 72 DEQTTVFFDKNIQGLAGRAILGKTSIKIWNTRRDSALYRCB--VVARNDKKEIDEI 129
DQ 153 -----HKGRDVLKKDV--RFIVLSNNVYLRGIKKTDEGTYRCGRILARG---EINFK 202
QY 130 VIELTVQKVPVTPVCR--VPKAVPVGMKATLHCQSEHGPRPHYSWYRNDVPLPTDSRAN 187
DQ 203 DIQVNVVPTVQARQSVNATANLQSVTLVC-DADGFPETMSWTXKGPEIENEER-D 260
QY 188 PFRNSSSHLNSSETGLVFTAVHKDDSGQYVCIASNDAGSARCEOEVEVY 238
DQ 261 EKRSSV-----SDSEVTRNVDRKDEAEYVCIAENKAG-----EQDASIH 302

RESULT 11
NEO1_RAT
ID NEO1_RAT STANDARD; PRT; 1377 AA.
AC P97603;
TX TISSUE=Brain;
RX MEDLINE=97015074; PubMed=8861902;
RA Keino-Masu K., Masu M., Hink L., Leonardo E.D., Chan S.S.-Y.,
RA Culotti J.G., Tessier-Lavigne M.;
RT "Deleted in Colorectal Cancer (DCC) encodes a netrin receptor.";
RL Cell 87:175-185(1996).
CC -!- FUNCTION: MAY BE INVOLVED AS A REGULATORY PROTEIN IN THE
CC TRANSITION OF UNDIFFERENTIATED PROLIFERATING CELLS TO THEIR
CC DIFFERENTIATED STATE. MAY ALSO FUNCTION AS A CELL ADHESION
CC MOLECULE IN A BROAD SPECTRUM OF EMBRYONIC AND ADULT TISSUES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 6 fibronectin type III domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U68726; AAC41100.1; -.
CC HSSP; P56276; 1TLK.
CC InterPro; IPR003961; FN III.
CC InterPro; IPR003962; FNIII subd.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_C2.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00041; fn3; 6.

DR PFam; PF00047; ig; 4.
DR PRINTS; PR00014; FNTYPEIII.
DR SMART; SM00060; FN3; 6.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS00835; IG-LIKE; 4.
KW Cell adhesion; Repeat; Signal; Transmembrane; Immunoglobulin domain;
KW Glycoprotein.
FT NON TER 1 1
FT SIGNAL <1 2
FT CHAIN 3 1377
FT CHAIN 3 1074
FT DOMAIN 1075 1095
FT TRANSMEM 1096 1377
FT DOMAIN 21 114
FT DOMAIN 121 206
FT DOMAIN 198 305
FT DOMAIN 310 395
FT DOMAIN 405 502
FT DOMAIN 505 598
FT DOMAIN 599 698
FT DOMAIN 704 798
FT DOMAIN 819 919
FT DOMAIN 920 1021
FT DOMAIN 1087 1090
FT DISULFID 43 98
FT DISULFID 142 190
FT DISULFID 239 289
FT DISULFID 331 379
FT CARBOHYD 42 42
FT CARBOHYD 179 179
FT CARBOHYD 295 295
FT CARBOHYD 439 439
FT CARBOHYD 458 458
FT CARBOHYD 608 608
FT CARBOHYD 684 684
FT CARBOHYD 878 878
SQ SEQUENCE 1377 AA; 150637 MW; E514ED8AD1A63A9 CRC64;

Query Match 10.5%; Score 171; DB 1; Length 1377;
Best Local Similarity 31.0%; Pred. No. 2.2e-06;
Matches 61; Conservative 21; Mismatches 69; Indels 46; Gaps 9;

QY 47 EVELSCITDSQTSQSDPRIWKIODEQTTVFFDKNIQGLAGRAEILGKTSKWNVT 106
DQ 233 QSAVLPCVA--SGLPAPVIRWKNEDVLT-----ESSGRLLALLAGSLEISDVT 280
QY 107 RDSALYRCBVARNDKKEIDEIVIELTVQKVPVTPVCRVPAVGMKATLHCQES--- 163
DQ 281 EDDAGTYFC--VADNGNKTI-EAQALTVQVPPF-----LKQPAIYARESMDI 327
QY 164 -----EGHPRPHYSWYRN-DVPLPTDSRANPRFNSSSHLNSSETGLVFTAVHKDDSGQ 216
DQ 328 VFEVTVGKPAPTVKNVNGDVIPSD-----YFKIVKEH-----NLQVLGLVKSDEGF 376
QY 217 YYCIASNDAGSARCEEQ 233
DQ 377 YQCIANDVGNQAQAQ 393

RESULT 12
NEO1_RAT
ID NEO1_RAT STANDARD; PRT; 858 AA.
AC P13596;
TX 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neural cell adhesion molecule 1, 140 kDa isoform precursor (N-CAM 140)
DE (NCAM-140).
GN NCAM1 OR NCAM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
```

FT	CARBOHYD	316	316	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	348	348	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	434	434	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	460	460	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	489	489	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SQ	SEQUENCE	858 AA;	94658 MW;	EALIA06A4EA0550F6	CRC64;

Query Match 10.4%; Score 170.5; DB 1; Length 858;  
 Best Local Similarity 26.3%; Pred. No. 1.4e-06;  
 Matches 61; Conservative 41; Mismatches 85; Indels 45; Gaps 12;

Qy	27	CLIGA-----VNLIKSNR-----TPVVOEFESVELSCIITDSOTS-DPRIEWKKIQ	71
Db	96	CVVTAEDGTQSEATVNVKIFQKLMFNKAPTPOQFKGEDAVIVCDVVSSUPPIIWK---	152
Qy	72	DEQTTVFFDNKIQGLAGRAELIGKTSKIMVNRDRSDALYRCE--VVARNDKRKEIDEI	129
Db	153	-----HKGRDVIILKDV--RFIVLSNLYIQIRGIKKTDEGTYCEGRILARG---	EINFK 202
Qy	130	VIELTVQVKPTVPCR--VPKAVPVGKMATLHCQSEGHPRPHYSWYRNDVPLPTDSRAN	187
Db	203	DIQIVNVPTVQARQISVNAVATNLGQSVTLVC-DADGFPPTMSWTKDGEPIENES---	258
Qy	188	PRFRNSSSHL-NSETGLVETAHVHKDDSGOYYCIASNDAGSARCEBQEMEVY	238
Db	259	---EDDEKHIFSDSDSELTRNVYDKNDEAFYGVCAENKAG-----EQDASIH	302

RESULT 13

TPPD	HUMAN	STANDARD;	PRT;	1912 AA.	
ID	PTPD	HUMAN	STANDARD;	PRT;	1912 AA.
AC	P23468;				
DT	01-NOV-1991	(Rel. 20, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	Protein-tyrosine phosphatase delta precursor (EC 3.1.3.48) (R-PTP-				
DE	delta).				
GN	PTPRD.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxId=9606;				
RN	[1]				
RN	SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-1178.				
RX	MEDLINE=95204468; PubMed=7896816;				
RA	Pulido R., Krueger N.X., Serra-Pages C., Saito H., Streuli M.;				
RT	"Molecular characterization of the human transmembrane protein-				
RT	tyrosine phosphatase delta. Evidence for tissue-specific expression of				
RT	alternative human transmembrane protein-tyrosine phosphatase delta				
RT	isoforms.";				
RL	J. Biol. Chem. 270:6722-6728(1995).				
RN	[2]				
RP	SEQUENCE OF 390-1912 FROM N.A.				
RC	TISSUE=Placenta;				
RX	MEDLINE=91006018; PubMed=2170109;				
RA	Krueger N.X., Streuli M., Saito H.;				
RT	"Structural diversity and evolution of human receptor-like protein				
RT	tyrosine phosphatases.";				
RL	EMBO J. 9:3241-3252(1990).				
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein				
CC	tyrosine + phosphate.				
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.				
CC	-1- ALTERNATIVE PRODUCTS:				
CC	Event=Alternative splicing; Named isoforms=3;				
CC	Comment=Additional isoforms seem to exist;				
CC	Name=1;				
CC	Isoids=P23468-1; Sequence=Displayed;				
CC	Name=2; Synonyms=Kidney;				
CC	Isoid=P23468-2; Sequence=VSP_005147, VSP_005148, VSP_005149;				
CC	Name=3; Synonyms=Fetal brain;				
CC	Isoid=P23468-3; Sequence=VSP_005150;				
CC	-1- PTM: A CLEAVAGE OCCURS THAT SEPARATES THE EXTRACELLULAR DOMAIN				
CC	FROM THE TRANSMEMBRANE SEGMENT				



```

DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; fn3; 2.
DR Pfam; PF00047; ig; 5.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00408; IGC2; 5.
DR PROSITE; PS50835; IG_LIKE; 5.
KW Cell adhesion; Transmembrane; Glycoprotein; Repeat;
KW Immunoglobulin domain; Signal.
FT SIGNAL 1 19
FT CHAIN 20 837
FT DOMAIN 20 697
FT TRANSMEM 698 718
FT DOMAIN 719 837
FT DOMAIN 21 108
FT DOMAIN 113 202
FT DOMAIN 208 297
FT DOMAIN 302 396
FT DOMAIN 401 491
FT DOMAIN 482 581
FT DOMAIN 594 678
FT DISULFID 42 93
FT DISULFID 136 186
FT DISULFID 232 282
FT DISULFID 322 380
FT DISULFID 422 475
FT CARBOHYD 177 217
FT CARBOHYD 219 309
FT CARBOHYD 406 406
FT CARBOHYD 419 419
FT CARBOHYD 445 445
FT CARBOHYD 474 474
FT CARBOHYD 562 562
SQ SEQUENCE 837 AA; 52932 MW; C3D034106C5741C1 CRC64;

Query Match 10.1%; Score 165; DB 1; Length 837;
Best Local Similarity 24.3%; Pred. No. 3.9e-06;
Matches 72; Conservative 31; Mismatches 105; Indels 88; Gaps 15;

QY 44 QEF---ESVELSCITDQSDPRLEWKIKDEQTYTFVFNKIOGDLAAGRAEILKGTSL 100
Db 124 QEFKQGEDAEVVCRRVSSPA--PAVSWLYHNEVTT--ISDN-----RLAMLANNL 171
QY 101 KINVTTRDSALYRCE--VVARNDKDEIDEIVIELTVQVKPTVPCVRPKAPVP----- 152
Db 172 QILNINKSDEGIYRCGEGRVEARGE-----IDFRDIIVVNPVPPAISMPPQKSFN 219
QY 153 -----GKMATLHCQSEGHPRPHYSWYRNDVPLPTDSRANPRFRNSSHLNSETGLTVFT 207
Db 220 ATAERGEEMTSCRAS-GSPEPAISFRNGKLE-----ENEKYLKGSNTELTIVR 269
QY 208 AVHKDDSGQYCIANDGASRCFQEMVEVDNLNGIIGVVLVAVLAITLIGICAY 267
Db 270 NIINSDGGPVCRATNKAG-----EDEKQAF-----LQVFQVPHIQLKNETTY 313
QY 268 RRGYFINNKQGESYKNQKPK-----DGUNYIRTEDEGDFR-----HKSS 307
Db 314 ENQGVTLVCDAB-----GEPPEITWKRAVDGFTTEGDKSPDGRKIEVKGQHSS 363

RESULT 15
PTK7 CHICK
ID PTK7 CHICK STANDARD; PRT; 1051 AA.
AC Q91048;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase-like 7 precursor (Kinase like protein).
GN PTK7 OR KLG.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=Embryonic brain;
RA MEDLINE=91271300; PubMed=1711213;
RX Chou Y.-H., Hayman M.J.;
RT "Characterization of a member of the immunoglobulin gene superfamily
RT that possibly represents an additional class of growth factor
RT receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:4897-4901(1991).
CC -!- FUNCTION: MAY FUNCTION AS A CELL ADHESION MOLECULE. LACKS PROBABLY
CC THE CATALYTIC ACTIVITY OF TYROSINE KINASE.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BONE MARROW, SPLEEN, BURSA,
CC THYMUS AND BRAIN. WEAKLY EXPRESSED IN FIBROBLASTS. ALSO EXPRESSED
CC IN EMBRYONIC LIVER.
CC -!- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
CC RECEPTOR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M63437; AAA48933.1; -.
CC PIR; A39712; A39712.
CC HSSP; P12931; 1FMK.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR00719; Prot_kinase.
CC InterPro; IPR001245; Tyr_kinase.
CC Pfam; PF00047; ig; 7.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC PRODOM; PD000001; Prot_kinase; 1.
CC SMART; SM00408; IGC2; 4.
CC SMART; SM00219; TyrKc; 1.
CC PROSITE; PS50835; IG_LIKE; 7.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Receptor; Transmembrane; Signal; Glycoprotein; Cell adhesion;
KW Immunoglobulin domain; Repeat.
FT SIGNAL 1 22
FT CHAIN 23 1051
FT DOMAIN 23 685
FT TRANSMEM 686 706
FT DOMAIN 707 1051
FT DOMAIN 23 105
FT DOMAIN 115 204
FT DOMAIN 213 298
FT DOMAIN 308 388
FT DOMAIN 393 472
FT DOMAIN 487 566
FT DOMAIN 573 661
FT DOMAIN 777 1048
FT DISULFID 40 88
FT DISULFID 137 187
FT DISULFID 234 282
FT DISULFID 326 372
FT DISULFID 414 462
FT DISULFID 505 551
FT DISULFID 594 645
FT CARBOHYD 103 103
FT CARBOHYD 202 202
FT CARBOHYD 255 255
FT CARBOHYD 264 264

```

FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 548 548 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 1051 AA; 116366 MW; 1752442AEA4CB702 CRC64;

Query Match 10.1%; Score 164.5; DB 1; Length 1051;  
Best Local Similarity 24.5%; Pred. No. 5.6e-06;  
Matches 66; Conservative 35; Mismatches 109; Indels 59; Gaps 11;

QY 40 TPVVQEFESVE-----LSCIITDSQSDPRIWKKIODEQTYVFFDNKIQGDLGRAE 93  
DB 486 TPPPQPLQCMENKVTWSCSATGRE--KPTIQWTKTD-----GSSLPSHVSHRAG 534  
QY 94 ILGKTSLKIMVNRDSDALYRCEVVARNDRKEIDEIVIELTVQVKPVTVCVRPKAVPVG 153  
DB 535 I-----LSFHKVSRSDSGNYTC--IASNSPQGEIRATVQLVVAVVVTFKLEPEPTTVYQG 587  
QY 154 KWATLHCOESEGHPRPHYSWYRNDVPLPTDGRANPRFRNSSSHLNSSETGLVFTAVHKDD 213  
DB 588 HTAMFQCQ-AEGDPVPHIQWKGDKIL-DPSKLLPRIQIMPN-----GSLVIYDVTTED 639  
QY 214 SGQYYCIASNDAG-----SARCEQE-----MEVVDLNIIGGIIGGVIV 251  
DB 640 SGKYTCTIAGNSCNIKHREAFLYVVDKPAAEDEGPSSHTPYKMIQTIGLSVGAAYAYIII 699  
QY 252 VLAVLALITLGICCAAYRRGYFINNKQDGE 280  
DB 700 VLGLMF-----YCKKRRKAKRLKKHPGE 723

Search completed: December 15, 2003, 14:52:33  
Job time : 20 secs